

# MAMADOU BALDE

AI Engineer | Data Scientist



## PROFIL

Data Scientiste avec 3+ ans d'expérience, spécialisé en NLP, LLMs et MLOps. Expert en conception de solutions IA déployées dans des environnements CI/CD. J'ai automatisé des pipelines ETL, construit des modèles prédictifs pour la 5G, et développé des chatbots intelligents.

## COORDONNÉES

Paris, Île-de-France, France

+33753266807

[baldeba772@gmail.com](mailto:baldeba772@gmail.com)

[LinkedIn](#)

[GitHub](#)  
[Hugging Face](#)

## COMPÉTENCES

- *Langages de Programmation :*
  - Python, C, SQL, Scala, R, SAS
- *Machine Learning & IA :*
  - Scikit-learn, TensorFlow, PyTorch, Transformers, LLM, RAG, LangChain
  - NLP, Computer Vision, Deep Learning, Embedding, Vector Store
  - Generative AI, LLM Framework, Fine-tuning
- *Visualisation & Business Intelligence :*
  - Power BI, Tableau, Looker Studio, Excel/VBA
- *Bases de Données & Stockage :*
  - PostgreSQL, MySQL, MongoDB, Neo4j, Microsoft SQL Server (SSAS, SSIS, SSRS)
- *Big Data & Cloud Computing :*
  - Apache Spark, Hadoop, HDFS, Databricks
  - AWS (SageMaker, EC2, S3, Lambda, ECS, ECR)
  - Azure (ML, Data Factory, Synapse Analytics)
  - Google Cloud Platform (BigQuery, Vertex AI, Looker Studio)
- *DevOps & MLOps :*
  - Docker, Kubernetes, Jenkins, Git/GitHub, Gerrit, CI/CD, MLFlow
- *Gestion de Projet :*
  - Agile (Scrum), Jira, Confluence, Project Management

## LANGUES

- **Anglais** : Courant (Certification Cambridge)
- **Français** : Langue maternelle

## EXPÉRIENCE PROFESSIONNELLE

### AI Engineer | Data Scientist, 09/2022 - 11/2024

Nokia Corporation – Massy Palaiseau

- Développé des modèles prédictifs (régression, classification) pour améliorer la fiabilité des builds CI 5G.
- Conception de dashboards Power BI pour le suivi des indicateurs clés (temps de build, taux de succès).
- Automatisé les pipelines ETL avec Python, SQL et Spark.
- Déploiement de solutions ML dans des containers Docker orchestrés avec Kubernetes.
- Piloté un projet stratégique pour le suivi automatisé de l'évolution du code (Python + Spark + Power BI).
- Travail en environnement Agile / international (squad CI\_Metrics, Tribes MN RAN).

### Assistant Data Scientiste | Analyste, 02/2021 - 04/2021

Michelin – Clermont-Ferrand

- Collaboré avec l'équipe R&D Michelin sur un projet transverse alliant data science et innovation automobile
- Analysé les comportements de conduite en extrayant des critères clés (accélération, freinage, trajectoires) à partir de données télémétriques
- Développé des modèles de classification supervisée pour optimiser la reconnaissance véhiculaire via l'analyse des comportements de conduite

### Stagiaire Data Scientiste | Analyste, 04/2020 - 08/2020

Laboratoire Mathématiques de Blaise Pascal - Aubière, 63170,

- Analyse statistique des taux de criminalité en France par département
- Clustering des données sociodémographiques pour détection de patterns
- Utilisation R, SAS et SQL pour modélisation et exploration des données.

## Projet personnel

### Projet NLP / LLM | RAG - Chatbot Médical IA

Développement d'un chatbot médical intelligent utilisant Retrieval-Augmented Generation (RAG) :

- Implémenté une architecture RAG avec LangChain pour interroger une base de connaissances médicales PDF
- Intégré le modèle avec embeddings sentence-transformers pour la compréhension contextuelle
- Utilisé FAISS pour l'indexation vectorielle et la recherche sémantique dans les documents médicaux
- Déployé sur HuggingFace Spaces pour un accès public : [Medical Chatbot Space](#)
- Technologies : Python, LangChain, Transformers, FAISS, Gradio, HuggingFace

## FORMATION

### MSc Data Science & AI, Data ScienceTech Institute, Nice 09/2021 -11/2023

- Spécialisation: MLOps, IA, Big Data, DevOps
- Certifications: Neo4j, AWS Certified Solutions Architect –Associate

### Master Mathématiques Appliquées & Statistiques, Université Clermont Auvergne (09/2019 – 04/2021)

### Licence Mathématique, Aix-Marseille-Université (09/2018 - 04/2019)

## CENTRES D'INTÉRÊT

- Sports : Musculation, Marathon, Football (Membre actif d'une équipe locale)
- Loisirs : Lecture (data science, IA, développement personnel), musique, manga